DTU Library

Insegt Fibre: a user-friendly software for individual fibre segmentation

Emerson, Monica Jane; Dahl, Anders Bjorholm; Conradsen, Knut; Dahl, Vedrana Andersen

Published in:
Proceedings of 22nd International Conference on Composites Materials

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

TWENTY-SECOND INTERNATIONAL CONFERENCE ON COMPOSITE MATERIALS
(ICCM22)

INSEGT FIBRE: A USER-FRIENDLY SOFTWARE FOR INDIVIDUAL FIBRE SEGMENTATION

Monica J. Emerson[1*], Anders B. Dahl[1], Knut Conradsen[1] and Vedrana A. Dahl[1]
[1] Department of Applied Mathematics and Computer Science, Technical University of Denmark, Kongens Lyngby, Denmark.
* Corresponding author (monj@dtu.dk)

Keywords: X-ray computed tomography, fibre detection, fibre tracking, microstructural characterisation, fibre orientations

ABSTRACT

Insegt Fibre is a software toolbox for volumetric fibre segmentation. The toolbox comes with scripts to detect the centres of individual fibres in 2D and 3D from tomograms acquired through X-ray imaging, and a graphical user interface to verify the accuracy of the resulting 3D tracks. In addition, there is a script to characterise fibre orientations in 3D and a script to match corresponding fibres across a 4D time-lapse sequence, which enables the characterisation of composite micro-structural changes of individual fibres. *Insegt Fibre* is based on a segmentation method that uses a dictionary of image patches which has been trained to model the patterns/features that are repeated in the image data at a certain scale defined by the patch size. Thus, the dictionary-learning segmentation algorithm has proven highly successful in modelling fibres that have regular cross-sections, commonly found in fibre reinforced composite materials. The algorithm is robust to noise and artefacts in the data and therefore excels in measuring fibre geometry from a variety of scan qualities, even when fibres are densely packed and the boundaries between them are unclear. *Insegt Fibre* is simple to use, it comes with a manual and it requires minimal input from the user. Besides presenting the recent user-friendly version of this robust method for measuring the 3D tracks of fibres from X-ray CT data, the paper gives an overview of the possibilities that the method gives with regards to characterisation of composite micro-structure and fibre behaviour under load. Thanks to the precision to which fibre geometry can be characterised with this method, it is now possible to follow how each individual fibre changes across data-sets acquired under progressive loading conditions. All in all, *Insegt Fibre* makes image-based characterisation of fibrous materials simpler, more accessible and applicable to a broader range of studies.

1 MOTIVATION AND SIGNIFICANCE

Fibre-reinforced composites are employed in a wide range of technologies for their high strength and stiffness relative to their weight. As the strength and stiffness properties of a composite are determined by its internal three-dimensional (3D) structure, X-ray computed tomography (CT) is becoming a popular tool for composite characterisation [1].

The recent advances in X-ray micro-CT have led to an increase in image quality for a given acquisition time and sample size or field of view (FoV). In particular, the improvement in spatial resolution has enabled the observation of individual fibres in scans capturing a larger FoV – large enough to be representative of the fibre structure. If interested in discovering biases in the process of fibre-bundle manufacturing, a representative FoV would contain a full bundle with tens of thousands of fibres.

Composite behaviour under load can be studied in-situ both at laboratory scanners and synchrotron beamlines. It is done by acquiring images while a sample is loaded, which makes it possible to monitor the geometric changes in the composite. Such data is 4D with three spatial and one temporal dimension. Thanks to the increase in temporal resolution obtained at the ultra-fast imaging set-ups found at some synchrotron beamlines, even the very small but sudden microstructural changes that lead to composite damage can be captured. This includes the moments in which the damage mechanism actually initiates, as well as the moments just before the sudden failure of the composite. Hence, the analysis of these 4D time-lapse data-sets will accelerate the understanding of the very complex damage mechanisms that affect fibre composites.

To measure the structural details of the vast amount of fibres contained in a bundle, or the very small changes across load steps in 4D data, quantitative image analysis is required. The first step is typically the detection of individual fibres using image segmentation. The accuracy of this step is crucial, as it defines the geometry of the material phases on which the subsequent analyses will be based, and ultimately the material characterisation.

Several methods have been proposed for detecting fibres in scans acquired with synchrotron light, e.g. in [2-3]. In both studies, individual fibres are segmented by setting a threshold on the intensities of the image. A segmentation based on intensity thresholding is however highly susceptible to noise and artefacts in the image, as well as to the contrast between the material phases.

For composites where fibres are densely packed, more robust image segmentation methods are necessary to handle a variety of image qualities. Methods that are relatively robust to noise and artefacts will facilitate the analysis of e.g. fast scans acquired while loading a composite in-situ, or scans acquired with relatively coarse pixel sizes (4 px/fibre) to image a larger (and thus more representative) sample region in a single scan. What is more, lower requirements to the scan quality translate into lower requirements for the X-ray set-up and shorter scan times. In other words, robust image segmentation algorithms will both broaden the application of composite imaging and make it accessible to a larger community of scientists.

More elaborate fibre segmentation algorithms have been proposed for lower quality scans of densely packed fibres in [4-6]. While Czabaj et al. [4] and Sencu et al. [5] developed complex tracking methods, our method described in Emerson et al. [6], has until now focused on the robustness, automation and speed of the fibre detection step. The dictionary-based segmentation method is able to analyse images of low resolution and with high noise levels where a good accuracy is obtained for a large range of parameter values, as demonstrated in [7]. In the template matching segmentation approach used by Czabaj, six fibre cross-sections are selected from the data manually to represent the appearance of the fibres. Instead, the probabilistic dictionary segmentation approach used here learns the representative patches from the data. The number of patches employed to model the data-set is much larger (in the order of hundreds) and can be increased to model a larger variety of fibre cross-sections. For example, depending on the local inclination of the fibres with respect to the scanning axes, the fibres may vary in size (diameter) and shape (e.g. ellipse axes ratio).

This paper presents *Insegt Fibre*, an interactive software tool for detection of individual fibres in 3D, that is now available online for free. The software is based on the latest version of the dictionary-based segmentation algorithm, first presented for 3D detection of individual fibres trajectories in [6].

2 SOFTWARE DESCRIPTION

Insegt Fibre is a software toolbox to measure individual fibres from X-ray CT scans of fibre-rich composites. Section 2.1. gives an overview of the methodology and algorithms on which the toolbox is based. Section 2.2 describes the input that is necessary from the user. To end with, Section 2.3 details the material that comes inside the *Insegt Fibre* toolbox.

2.1 Fibre detection algorithm

The aim of this section is to describe the changes in methodology since [6]. To understand the changes, a short overview of the method is provided, and the relevant dictionary concepts are presented. For more information, the reader is referred to [6].

2.1.1 The pipeline for measuring the 3D tracks of individual fibres

Insegt Fibre is based on a method that detects the 3D centre lines of individual fibres in two steps, these are shown in Fig. 1. The first step concerns the detection of fibre cross-sections in 2D images. In the case of unidirectional fibre-reinforced composites, the 3D scan can be sliced such that the resulting 2D images show fibre cross-sections of similar shape (circular) and size (diameter). The second step concerns tracking, i.e. matching centre points from one slice to the next, so as to define determine the 3D centre lines of the fibres contained in the composite volume.

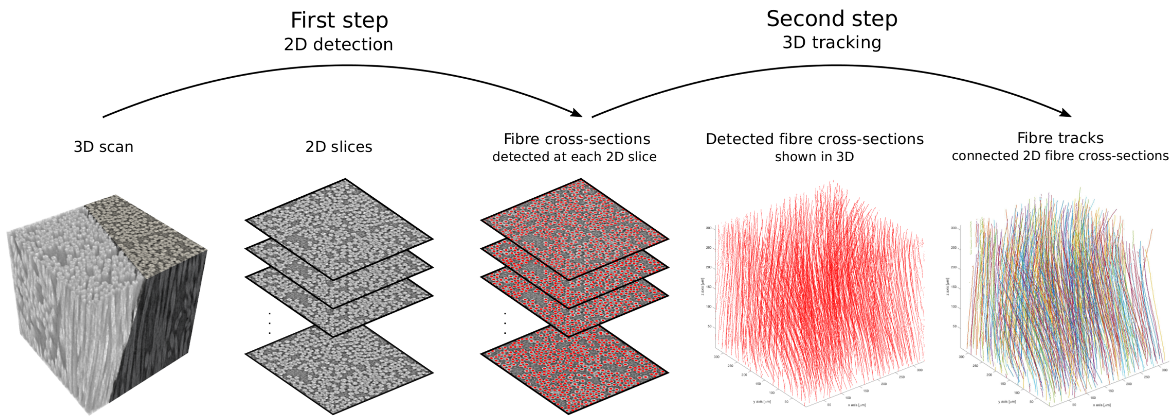


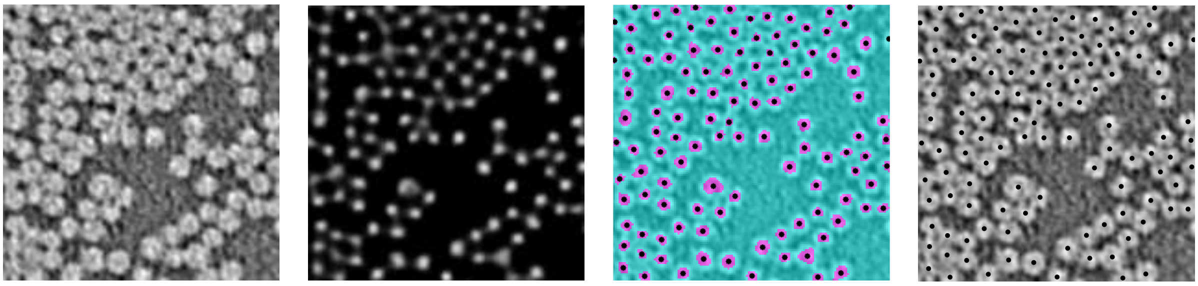
Figure 1: The *Insegt Fibre* pipeline. In the first step, fibre cross-sections are detected from 2D images. This process involves i) slicing the 3D scan in the direction orthogonal to the fibres and ii) applying the dictionary model learnt from the data to find the centres of the fibre cross-sections at each 2D slice. In the second step, the 3D fibre tracks are defined by connecting the fibre cross-sections that correspond to the same fibre. This step involves matching fibre detections along the third dimension, based on the cross-sectional distance between the points to be matched.

2.1.2 Detection of fibre cross-sections from 2D images

The first step concerns finding the coordinates of the fibre cross-section centres at every 2D slice. Fig. 2 illustrates this process. A dictionary model learnt from a training data-set is applied to an intensity image (see Fig. 2(a)) to obtain its probabilistic segmentation (see Fig. 2(b)). The probabilistic segmentation (or probability map) indicates how likely it is that a pixel belongs to the central region of a fibre. The segmentation of fibre-centre regions is obtained from the probability map by setting a threshold on the probability value. That is, all the pixels with a probability above the threshold value are labelled as belonging to the fibre-centre class. The fibre-centre class (*pink* in Fig. 2(c)) grants the segmentation of the individual fibres. The coordinates of the fibre centres (*black* dots in Fig. 1(c-d)) are computed as the centroids of the *pink* connected components.

The training data-set comprises a 2D image that has been annotated by the user, to indicate whether a pixel belongs to a fibre-centre region (i.e. is close to a fibre centre) or not. The dictionary of images patches is comprised of a dictionary of intensity patches, and a corresponding dictionary of label patches. In the fibre case, the labels indicate whether a pixel belongs to a fibre-centre region (i.e. is close to a fibre centre) or not. A dictionary of image patches is a representative and compact model of the features that are repeated in the training data-set at the patch scale. In the fibre case, the features that are repeated are the fibre cross-sections (similar in shape and diameter).

To obtain the probabilistic segmentation of an image, each intensity patch in the image is matched with an element in the dictionary of intensities using a metric for patch similarity (e.g. Euclidean distance between patch intensities). The corresponding element from the dictionary of labels is employed to build the probability map. The resulting probability value for each pixel will be the average of the label patches that correspond to the overlapping intensity patches covering that pixel. For more details, the reader is referred to [6].



(a) Intensity image (b) Probability map (c) Segmentation (d) Detected fibre centres

Figure 2: The steps in detecting fibre centres from a 2D cross-sectional slice. The image in (a) is a small region of a cross-sectional slice from the 3D scan. The dictionary-based probability map in (b) is segmented into the magenta regions shown in (c). The centroids of the segments are marked as black dots. These are shown on top of the original image in (d).

2.1.3 Building the dictionary model

The dictionary of intensities is built from the training data-set by clustering the training patches according to their intensity. The patches are square and their width, measured in pixels, is defined by the patch size M. The training patches are Ψ patches selected randomly from the training image. The approach for computing the elements in the dictionary of intensities has changed since [6]. In [6], the n elements in the dictionary of intensities are the cluster centres obtained via the weighted k-means algorithm. Now in the latest implementation, the n elements in the dictionary of intensities are the nodes of a tree built using hierarchical k-means clustering [8]. While in [6] there are as many dictionary elements as clusters in the k-means algorithm, i.e. $n = k$, now the number of elements is equal to the number of nodes in the tree, i.e. $n = \sum_{i=1}^{l} b^i$, where l is the number of layers and b the number of branches.

Both versions of the k-means algorithm (weighted and hierarchical) achieve clustering of training patches. Hierarchical is faster. The quality of clustering is slightly less for hierarchical, but our method proves to be extremely robust to the quality of clustering. In both versions, the dictionary of labels is built by averaging the label patches associated to the intensity patches in each cluster.

As the weights for the k-means clustering in [6] were defined by the labelling of the training image, the training image had to be fully labelled before computing the dictionary of intensities. In the latest version, the clustering is independent of the labelling, so the dictionary of intensities can be computed before the user provides the manual labelling. Once the dictionary of intensities is computed, it is possible to find the linear mapping that associates pixels in the image and the dictionary as shown in Fig. 3, because the dictionary elements are obtained from the image patches by summations and averages. This linear mapping means that information can be transformed from the image to the dictionary domain (or vice versa) with a matrix multiplication, which can be computed fast.

Thanks to the pre-computed linear mapping between dictionary and intensity domains, it has been possible to obtain real-time response to user input. Therefore, we developed a graphical user interface (GUI) for learning the dictionary of labels interactively. Every time the user places a label on the training image, the dictionary is updated and used to segment the full training image, which is shown immediately to the user. In this way, the user can stop labelling pixels when satisfied with the segmentation of the full training image. This partial interactive labelling approach, where the user is able to see the effect of the manual annotations immediately, has significantly reduced the amount of manual input required to learn an adequate dictionary model.

As mentioned above, the manual annotation provided by the user is propagated to the dictionary and back to the image domain via two matrix multiplications. There are several methods for label propagation and these are detailed in the following. In all methods, the starting point is the label image defined by the partial annotations. The pixels in the label image take values of ‘1’ or ‘0’ where annotations have been provided and ‘1/2’ elsewhere. While the one-step method (named *distributed*) stops once the label image has been propagated to the dictionary and back to the image domain once, the two-step methods propagate the information once again through the system, so that the partial annotations get distributed to as many pixels as possible in both the training image and the dictionary. The input label image for the second propagation step varies according to the propagation method. For the two-step thresholded method (*two_max*), the label image is the thresholded probability map. It is the same for the two-step thresholded method with overwrite (*two_max_over*), except for the pixels that had been annotated by the user, those take the manual labelling. For the two-step continuous (*two_cont*) and the two-step continuous with overwrite (*two_cont_over*) it is the probability map, overwritten with the manual labelling in the later.

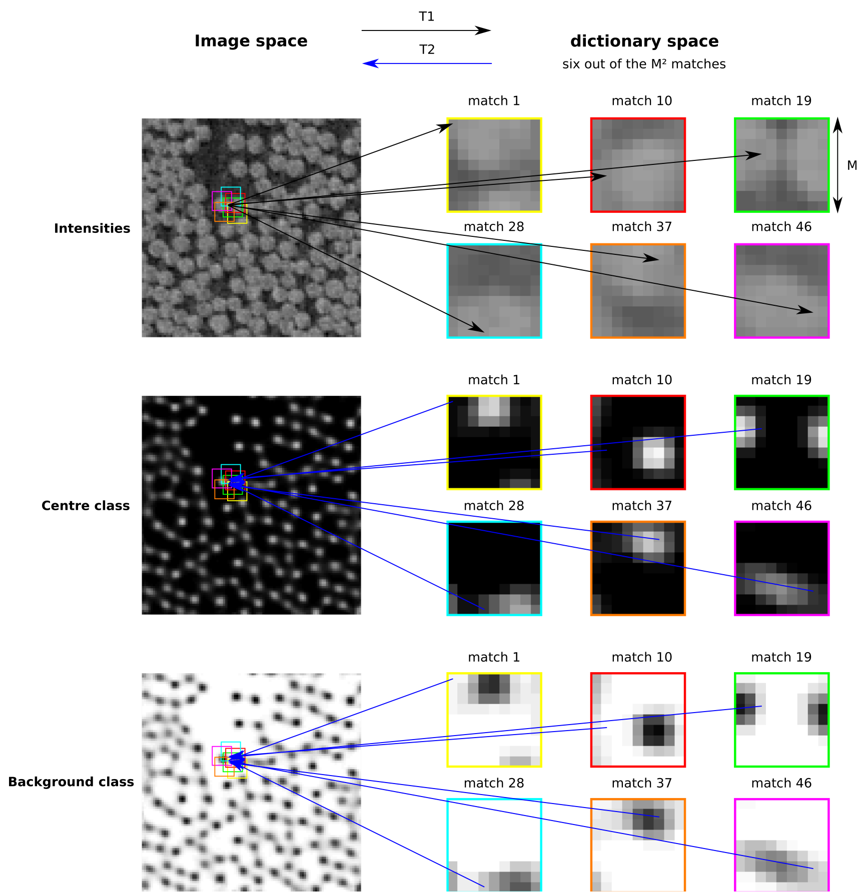


Figure 3: Graph description of the relationship between image and dictionary spaces. Each image pixel is connected to the M^2 dictionary elements that have been matched to the M^2 overlapping patches covering that pixel. The link is to a specific pixel of the dictionary element. The location of this pixel is determined by the position that the image pixel has inside the overlapping image patch.

2.1.4 Improvements in the dictionary-based segmentation method for 2D images

Our method for segmenting images based on a dictionary of image patches is under continuous development. We hereby sum up the improvements in the dictionary learning phase and the segmentation phase.

In the dictionary learning-phase the latest implementation reduces the computational time spent in building the dictionary, and more importantly the amount of manual labelling required from the user to learn the dictionary. With the latest interactive partial labelling approach, the amount of user input is minimal and adaptable to the data-set at hand. In the segmentation phase, the computational time required for segmenting a complete scan using the learnt dictionary model has been drastically reduced.

The task of segmenting a complete scan can be highly time-consuming because it requires finding the closest match in the dictionary of intensities for every overlapping patch in the volume.

The biggest reduction in computational time comes in the segmentation phase, where the new tree-based approach speeds up the process of finding a match in the dictionary. A slight speed-up in this task leads to a significant speed up in the segmentation phase. While the dictionary is only built once, the task of finding a match in the dictionary is carried for each pixel in the volume, except for those that are too close to the edges. A pixel is too close to the edge of the 3D scan when an image patch cannot be centred over that pixel. Hence, assuming that the patch size M is much smaller than the dimensions of the volume, only a small number of pixels will be to close to the edge, and the operation of finding a dictionary match will be performed for almost every pixel. For a typical 3D scan of size 1000^3 pixels, the dictionary is searched approximately 10^9 times. While it took 3 hours to segment a volume of 1000^3 pixels with a Macbook Pro 2,2 GHz Intel Core i7 and a dictionary of $n = 250$ elements, it now takes under 50 minutes. It is of course possible to speed up both implementations with a factor of p, by processing p slices in parallel on p cores.

2.1.4 Fibre tracking

In the case of unidirectional fibre composites, as our robust segmentation method provides accurate fibre detections, the tracking is simply based on the cross-sectional distance between points in contiguous slices. In [6] fibre detections were matched in a unidirectional manner. This means that a detection in slice t was connected to the closest point in slice $\mathrm{t+1}$, as long as they were closer than a certain distance. After observing that two detections in slice t had been connected to the same point in slice $\mathrm{t+1}$, the tracking approach was updated. In the latest version the tracking is bidirectional, meaning that fibre detections ought to be each other's nearest neighbour in order to be connected. That is, the detection from slice t should also be the closest to the detection from slice $\mathrm{t+1}$.

2.2 User input

This section focuses on the input required from the user of *Insegt Fibre* and gives some tips to aid the user in providing adequate inputs. The first inputs are required to learn the dictionary of image patches. The input required from the user at this stage comprises the dictionary parameters and the training data-set from which the dictionary model will be learnt. The next inputs are required to obtain the 3D fibre tracks with the learnt dictionary model. First, a threshold is necessary to obtain the 2D fibre centre detections from the dictionary-based probabilistic segmentation of the 2D slices. Then, when the 2D detections are connected via tracking to form 3D fibre centre lines, the user can set a distance. Fibre detections will be connected from one slice to the next as long as they are closer than that distance.

2.2.1 Dictionary parameters

The parameters that define the dictionary characteristics are meant to be simple and meaningful, and most are actually pre-set to adequate default values. There is really just one parameter that must be set by the user, and that is the patch size M. As the patches are square and, for convenience, placed over a central pixel, the width of a patch (or patch size) will be an odd integer. It was mentioned previously that the dictionary is a good model of the features that are repeated at the patch scale. To model fibre cross-sections, the patch should cover a little more than a full fibre diameter. Thus, the patch size M is computed as the odd integer that is closest to the ratio between the fibre diameter and the pixel size of the scan. Ideal patch sizes are 7, 9, 11 or 13. If the patch size is 5 or smaller, it is highly recommendable to upscale the slices in the scan such that the new pixel size grants a patch size larger than 7. This is done to avoid a loss of precision in the centre coordinates of the fibre cross-sections. For patches of size larger than 15, the user could consider downscaling the slices to reduce memory usage.

The user can choose to set other dictionary parameters in some cases. For example, when the variety in the data-set is large, i.e. fibre cross-sections vary significantly in shape in diameter because of e.g. high misorientation in the sample, increasing the number of elements in the dictionary is recommended.

Bear in mind that the dictionary is learnt by clustering a subset of Ψ training patches, so there should be many more training patches than dictionary elements. For the default number of dictionary elements $n = 363$, resulting from a branching factor $k = 3$ and a $l = 5$ layers, $\Psi = 15000$ training patches are used. Thus, if increasing the number of dictionary elements, remember to increase the number of training patches, such that it is in between one and two orders of magnitude larger. And of course, as the training patches are randomly selected from the overlapping patches extracted from the training image, the number of training patches should not be larger than the number of patches in the training image. The number of overlapping patches in a 2D image is approximately equal to the number of pixels in the image. Thus, for a typical training image of size 200×200 pixels, $\Psi \leq 40000$.

2.2.2 Training data-set

Insegt Fibre will ask the user to select the 2D image region from which the dictionary model will be learnt. To select the training region, the user will first be asked to choose a cross-sectional slice of the tomogram and then to select a smaller region from the chosen slice. There is the option of selecting this region by either defining the start and end coordinates of the *x* and *y* axes, or interactively, by placing a rectangle over the slice. To obtain a dictionary that models the sample adequately, the training region should be representative of the sample. A training image containing between 100 and 1000 fibres is usually large enough to capture the variation in fibre-packing density and fibre cross-section (shape and diameter), and is still of a good size for the annotation process explained in the following.

The amount of manual annotation required to learn the dictionary is typically very little. Whereas in [6] the user had to label the full training image, the dictionary of labels can now be computed from a partial annotation of the training image. The purpose of providing annotations, indicating whether a pixel belongs to the central region of a fibres cross-section or not, is to learn the dictionary of intensities, so that the dictionary not only represents the patterns in our image, but can also identify the pixels that are close to a fibre centre.

To facilitate the manual annotation process, the software toolbox comes with a graphical user interface. As shown in Fig. 4, the GUI displays two images of the same composite region, where pixels are labelled in *pink* if they are close to fibre centres and *blue* otherwise. Every time the user places new markings on the left-hand side image, the segmentation of fibre cross-sections on the right is updated. As soon as the user is satisfied with the segmentation of the fibre centres, the dictionary is ready to process the full volume. If the training image is too large, it will be time consuming to verify that all fibre centres are captured on the right. That is why we recommend a region in between 100 and 1000 fibres, large enough to capture the variety in the sample, but small enough to verify at a glance.

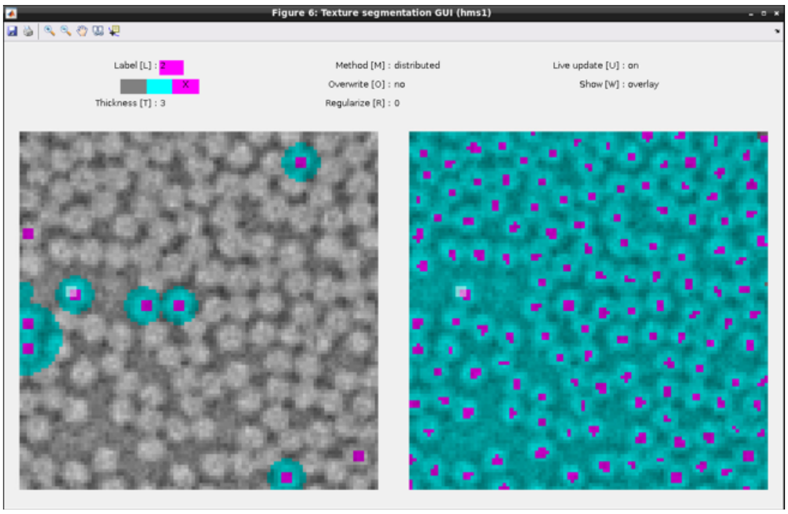


Figure 4: Graphical user interface (GUI) for training the dictionary model that will detect fibre cross-sections in the full volume. On the left the manual annotations, and on the right the detected fibre centres.

Some functionality of the user interface is controlled by keys. The key ‘w’ is used to switch the content displayed on the right-hand side: 1) ‘segmentation’ (obtained after thresholding the probability map for the centre class with a value of 0.5), 2) ‘overlay’ (segmentation over intensity image) or 3-4) probability maps for centre or non-centre class respectively. The key ‘l’ switches labels and ‘t’ changes the thickness of the brush. The key ‘m’ is used to explore different methods for propagating labels. As to the other keys, the default values are recommended.

2.2.3 Input for the detection of fibre centres in 2D and 3D

The centre detections are obtained from the hard segmentation, which determines which pixels belong to a fibre-centre region, as is illustrated in Fig. 2. This segmentation is obtained from the probability map by setting a threshold on the probability value. The threshold value should be somewhere in between 0.4 – 0.6. It is recommended to start by using the default value of 0.5 and explore the resulting segmentation. Change the threshold to enlarge or reduce the centre regions to e.g. detect all fibres or separate merged fibres.

Once the fibre centres have been detected in the full stack of 2D cross-sectional slices, these are tracked to connect detections that correspond to the same fibre. Detections are matched when they are the closest neighbour to each other, but only as long as they are closer than a certain distance. This distance defines the maximum fibre displacement. In other words, the amount of cross-sectional displacement allowed between two points to be matched. By default, the maximum fibre displacement is set to 5 pixels, reduce this value to be more conservative in the tracking. Even for misalignment angles as high as 45°, the displacement from one slice to the next is just 1 pixel. However, it might be that for speed reasons the centres have just been detected in 1 out of 10 slices. In this case, the displacement of a fibre that is oriented at 45° with respect to the z scanning axis will be 10 pixels. As long as the maximum fibre displacement is smaller than half of the fibre diameter, there is not risk of connecting to surrounding fibres that are in contact. Typically, fibre diameters cover in between 4 to 16 pixels.

2.3 The example scripts and manual

Insegt Fibre comes with a manual, four example scripts written in Matlab® and example 2D and 3D data-sets. The user manual contains guided exercises to run the user through the scripts step by step, so that *Insegt Fibre* can be executed without prior programming experience. The first script focuses on the dictionary training step. It teaches the user how to set the patch size and annotate in the GUI to analyse four composite data-sets of varying complexity. The second script guides the user through the complete segmentation pipeline, to measure the 3D fibre centre lines from a full volume, as shown in Fig. 1. Download *Insegt Fibre* from Science Cases at http://qim.compute.dtu.dk.

2.3.1 *Insegt Fibre* for 2D data

The script `InsegtFibre_2D.m` gets the user familiarised with the input required to train the dictionary model (parameter dictionaries and training data-set) and detect the fibre centres from 2D images (threshold applied to the probabilistic segmentation). The first exercise of the manual describes the script step by step and guides the user in the analysis of four data-sets of varying quality. These data-sets are scans of unidirectional (UD) glass and carbon fibre reinforced polymers (GFRP and CFRP) acquired at either a synchrotron or a laboratory X-ray source. The 2D images used in this exercise are cross-sectional slices of the 3D CT scans. The 2D images are shown in Fig. 5, see Table 1 for the fibre diameter and pixel size. Additional information can be obtained in the articles where these material systems were investigated [6,8,9,10].

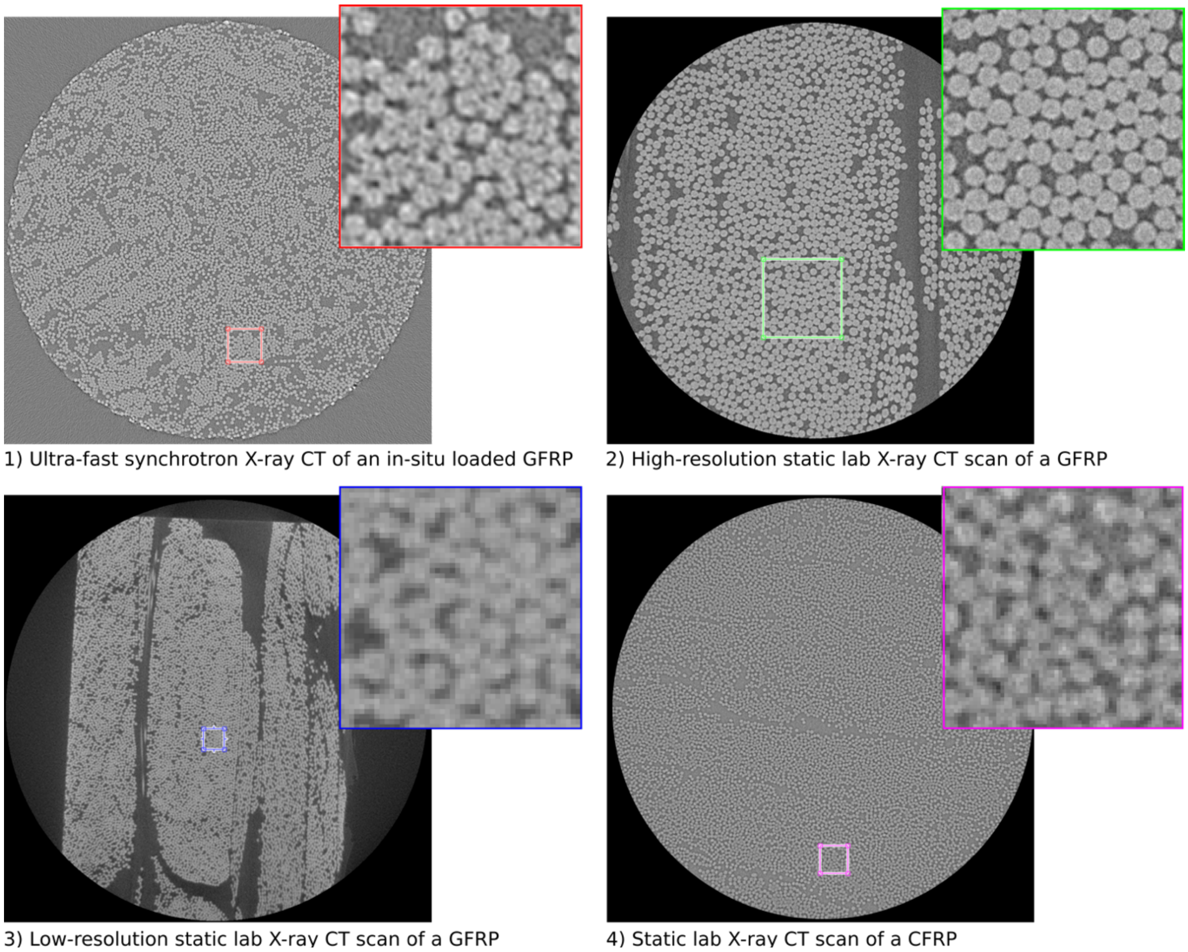


Figure 5: 2D cross-sectional slices from X-ray CT scans of unidirectional fibre-reinforced polymers. These scans of pixel size and fibre diameter given in Table 1 are used as examples for learning how to train a good dictionary model that will adequately segment the centres of fibre cross-section, so as to detect individual fibres from 2D images.

Data-set, scan type and composite material	Pixel size [µm]	Fibre diameter [µm]
1) Ultra-fast synchrotron CT of in-situ loaded GFRP [8]	1.10	12
2) High-resolution static lab X-ray CT of GFRP [9,10]	1.04	17
3) Low-resolution static lab X-ray CT of GFRP [9,10]	2.81	17
4) Static lab X-ray CT of CFRP [6]	0.96	7

Table 1: Pixel sizes and fibre diameters for the example data-sets used in *Insegt Fibre*.

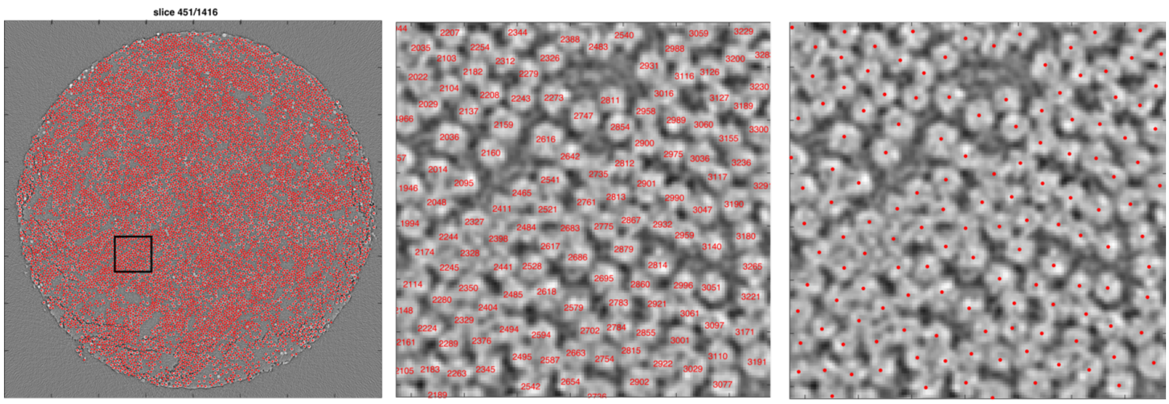
2.3.2 *Insegt Fibre* for 3D data

The script `InsegtFibre_3D.m` guides the user through the complete pipeline of *Insegt Fibre*. As shown in Fig. 1, the input to the pipeline is a 3D X-ray CT scan and the output is a set of 3D centre lines representing the tracks of the individual fibres. Again, there is an exercise in the manual to guide the user through the sequence of steps in the pipeline for detection of 3D fibre tracks. The example data-set employed here is data-set number 1 (see Fig. 5 and Table 1). The 3D image is part of a 4D time-lapse data-set acquired while compressing a UD GFRP progressively in the fibre direction. The 3D image was acquired under 0N load and can be downloaded from http://doi.org/10.5281/zenodo.2597498 [10], it is named `GFRP_Initial.zip`. More details about the data-set can be found in [8].

2.3.3 *Insegt Fibre* for validation of measured 3D tracks

The script `InsegtFibre_validate3DTracks.m` provides a qualitative way of assessing the accuracy of the measured fibre tracks with a GUI. The user selects the tracks and the analysed volume (scan and coordinates defining the specific region). The first time the script is run for a certain volume, the volume is saved as one only '*.tif' file, containing the stack of slices composing the analysed region of the scan. The next time the script is run for that same volume (specific region of a scan), one can simply load the saved '*.tif' file.

Fig. 6 displays the possibilities of the GUI. As can be seen, the tracks are placed over the CT slices, so as to get an insight into the number of detected fibres and the precision of the detected centre lines. The arrows are used to navigate along the depth of the volume. The arrows '↑' and '↓' are used to move one slice at a time whereas the '→' and '←' jump 10 slices at a time. The user ought to zoom in into a region to see which fibre the detection belongs to, and then navigate along the depth of the volume with the arrows to follow the fibre tracks and verify that the right detections have been connected. The key 'n' can be pressed while zoomed in to assess the precision of the centre placement.



a) Zoom out view　b) Zoom in view　c) Zoom in view after pressing 'n'

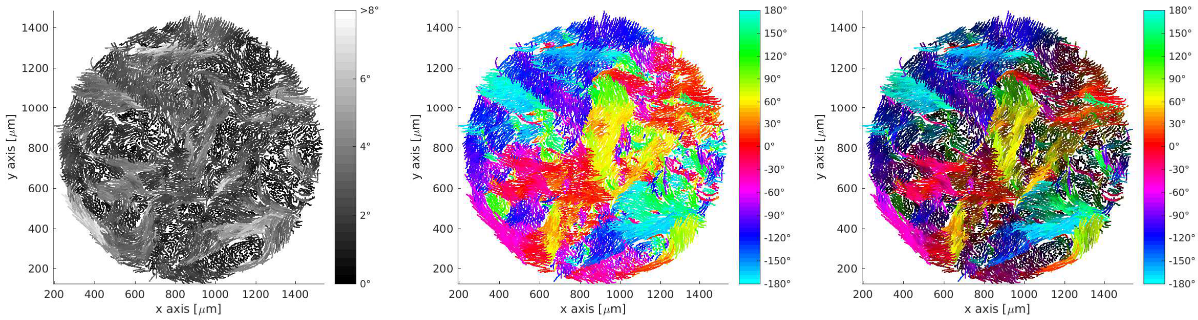
Figure 6: Graphical user interface for qualitative validation of the fibre tracks. (a) The default view when the GUI opens, a cross-sectional slice with the detections shown in red. (b) A closer view obtained after zooming in into the square marked over (a). Fibre numbers are displayed over the cross-sections, so that the accuracy of the tracking can be verified by slicing through the volume in the zoomed view. (c) The closer view after pressing the key 'n' to assess the precision in the placement of the centre detection.

2.3.4 *Insegt Fibre* for orientation characterisation

The script `InsegtFibre_orientations.m` computes for each fibre an inclination and an azimuth angle, based on the displacement between the start and end points of the fibre track. While the inclination indicates the magnitude of misalignment with respect to the z-axis, the azimuth is indicative of the direction in which a fibre is misaligned. The script will plot a histogram for the inclination angle and a histogram for the azimuth angle. In addition, it produces three plots of the fibre tracks coloured according to their orientation (see Fig. 7).

2.4 Accuracy and precision validation

We have used the GUI presented in Fig. 6 to validate the estimated fibre tracks qualitatively. The method's accuracy and precision has also been measured quantitatively. In [6], the accuracy was measured as the percentage of fibres detected in a test image region that contained around 1000 fibres. In the same study, we verified that the compression strength estimated from the measured fibre orientations was of the same order of magnitude as the compression strength measured via mechanical testing. In study [11], we demonstrated the precision of our centre detection and tracking. The study compared diameter estimations across two analysis methods (probabilistic dictionary segmentation and Hough transform) and six scans (varying in modality and resolution) of the same sample region.



a) Inclination b) Azimuth c) Azimuth and Inclination

Figure 7: Top views of measured fibre tracks, color-coded according to their orientation. Fibres appear as dots when they are aligned with the z-axis.

3 EXAMPLE APPLICATIONS AND IMPACT

The first application of *Insegt Fibre* is the characterisation of real composite microstructure through fibre orientations and diameters. *Insegt Fibre* was applied with success to study the spatial distribution of orientations in glass and carbon fibres [6] and diameters in full-bundle scans of glass fibres with just 4 px/fibre [10]. Besides getting insights into the variabilities introduced by the manufacturing process, material properties such as stiffness or strength can be computed analytically from the measured orientations and diameters. Based on the measurements provided by *Insegt Fibre*, the compression strength was estimated from the average fibre orientation in [6] and the transverse stiffness from the volume fraction given by the fibre diameters in [11]. For more precise estimates of the mechanical properties, micromechanical models can be created from the image-based fibre geometry. In [11], the transverse stiffness was obtained from a micromechanical model based on the geometry measured with *Insegt Fibre*.

The other main application of the pipeline is the study of fibre composite behaviour under realistic working conditions. By combining ultra-fast X-ray CT, in-situ loading of a sample and *Insegt Fibre*, it has been possible to accurately quantify changes in fibre trajectories fibre by fibre. Quantifying the changes in composite microstructure with a high degree of accuracy both in space, time and load can provide valuable insight into the behaviour of materials under real-life loading conditions, which will lead to a better understanding of the complex damage mechanisms that affect composite materials. In particular, *Insegt Fibre* has provided valuable insight with regards to the behaviour of a unidirectional fibre composite under compression loading in the fibre direction [8]. The study shows for the first time that fibres start to deflect at very low loads and in a direction that is related to the final damage.

4 CONCLUSIONS

Insegt Fibre is based on a simple, computationally efficient, while also accurate method to measure fibre geometry from tomograms acquired through X-ray imaging. The dictionary-based method is appropriate for fibres that have similar cross-sectional shape and diameter, because the dictionary model learns the patterns/features that are repeated in the data at the patch scale. By increasing the number of elements (image patches) in the dictionary, it is possible to model a larger variety of fibre cross-sections, e.g. elliptical fibres varying in diameter.

The strength of this methodology lies in its ability to measure individual fibres from volumetric scans where there is a high fibre volume fraction. Thanks to the method's robustness to noise and the fact that it can cope with low contrast between material phases, the image data can be acquired under shorter scan times and a broad range of fibre materials can be investigated. Additionally, the method can handle low spatial resolutions where fibre boundaries are hard to resolve, and fibres are challenging to separate individually. The method's capability of dealing with challenging scans has enabled the investigation of representative sample sizes from a single scan and time-lapse scans acquired while exposing the material to realistic loading conditions.

As demonstrated, the method has great potential in providing insights for understanding fibre microstructure, which can then be related to composite manufacturing and mechanical properties. *Insegt Fibre* can be freely downloaded, and we hope that this can become a useful tool for understanding the structural properties of fibre composites.

ACKNOWLEDGEMENTS

The paper has been made as part of QIM: The Centre for Quantification of Imaging Data from MAX IV supported by the Capital Region Denmark.

REFERENCES

[1] E. Marie and P. J. Withers, Quantitative X-ray tomography, *International Materials Reviews,* Vol. 59, no. 1, pp. 1–43, 2014.

[2] G. Requena, G. Fiedler, B. Seiser, P. Degischer, M. Di Michiel and T. Buslaps, 3D-Quantification of the distribution of continuous fibres in unidirectionally reinforced composites, *Composites Part A: Applied Science and Manufacturing*, Vol. 40, no. 2, pp. 152–163, 2009.

[3] G. Gaiselmann, I. Manke, W. Lehnert and V. Schmidt, Extraction of curved fibers from 3D data, *Image Analysis and Stereology*, Vol. 32, no. 1, pp. 57–63, 2013.

[4] M. W. Czabaj, M. L. Riccio and W. W. Whitacre, Numerical reconstruction of graphite/epoxy composite microstructure based on sub-micron resolution X-ray computed tomography, *Composites Science and Technology*, Vol. 105, pp. 174–182, 2014.

[5] R. M. Sencu, Z. Yang, Y. C. Wang, P. J. Withers, C. Rau, A. Parson and C. Soutis, Generation of micro-scale finite element models from synchrotron X-ray CT images for multidirectional carbon fibre reinforced composites, *Composites Part A Applied Science and Manufacturing*, Vol. 91, pp. 85–95, 2016.

[6] M. J. Emerson, K. M. Jespersen, A. B. Dahl, K. Conradsen and L. P. Mikkelsen, Individual fibre segmentation from 3D X-ray computed tomography for characterising the fibre orientation in unidirectional composite materials, *Composites Part A: Applied Science and Manufacturing*, Vol. 97, pp. 83-92, 2017.

[7] M. J. Emerson, K. M. Jespersen, P. S. Jørgensen, R. Larsen and A. B. Dahl, Dictionary based segmentation in volumes, *Proceedings of the 19th Scandinavian Conference on Image Analysis,* Springer, Cham, pp. 504-515, 2015.

[8] D. Nister and H. Stewenius, Scalable recognition with a vocabulary tree, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* Vol. 2, pp. 2161–2168, 2006.

[9] M. J. Emerson, Y. Wang, P. J. Withers, K. Conradsen, A. B. Dahl and V. A. Dahl, Quantifying fibre reorientation during axial compression of a composite through time-lapse X-ray imaging and individual fibre tracking, *Composites Science and Technology*, Vol. 168, pp. 47–54, 2018.

[10] M. J. Emerson, V. A. Dahl, K. Conradsen, L. P. Mikkelsen and A. B. Dahl, A multimodal data-set of a unidirectional glass fibre reinforced polymer composite, *Data in Brief,* Vol. 18, pp. 1388–1393, 2018.

[11] M. J. Emerson, V. A. Dahl, K. Conradsen, L. P. Mikkelsen and A. B. Dahl, Statistical validation of individual fibre segmentation from tomograms and microscopy, *Composites Science and Technology*, Vol. 160, pp. 208–215, 2018.

[12] L. P. Mikkelsen, M. J. Emerson, K. M. Jespersen, V. A. Dahl, K. Conradsen, A. B. Dahl, X-ray based micromechanical finite element modeling of composite materials, *29th Nordic Seminar on Computational Mechanics Nordic Seminar on Computational Mechanics*, 2016.